Análise de Regressão

Mariana Costa Freitas

2024-04-24

Resumo

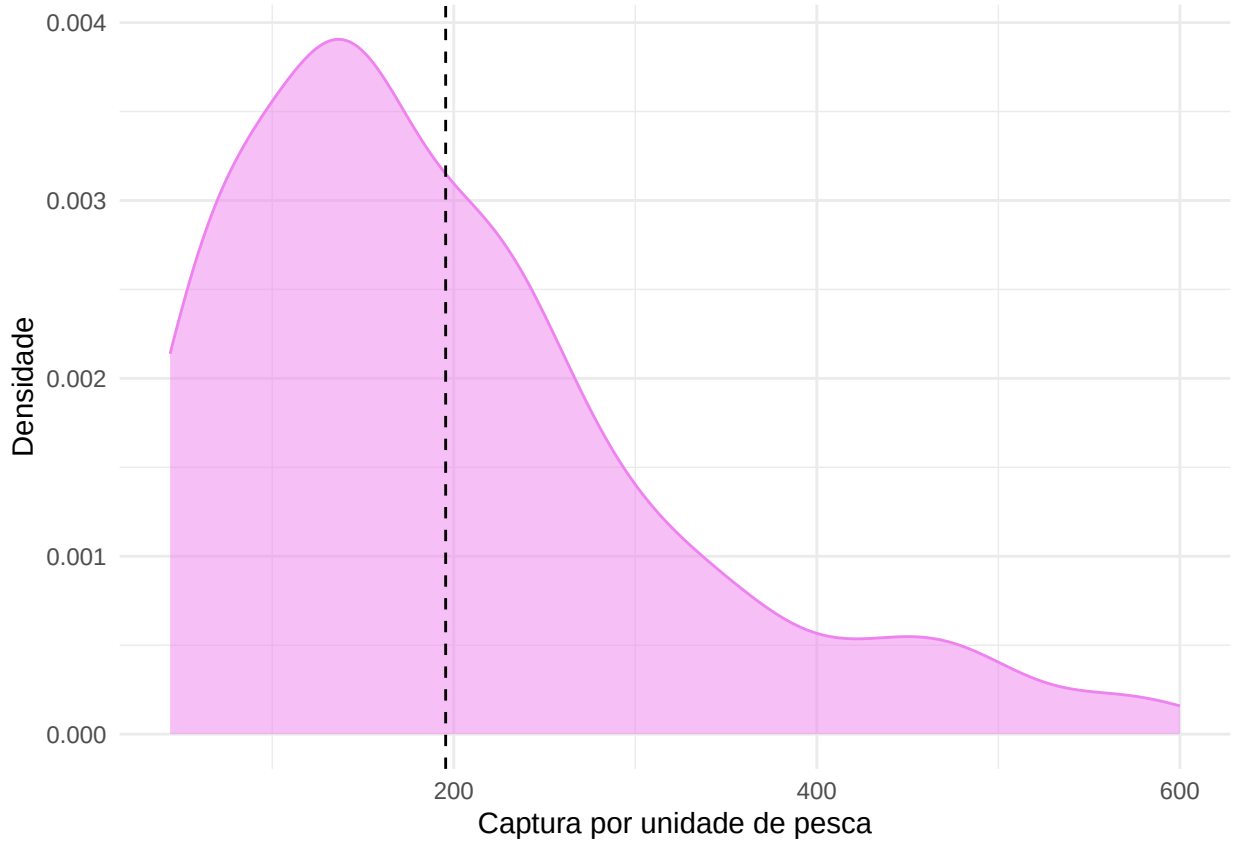
Introdução

Metodologia

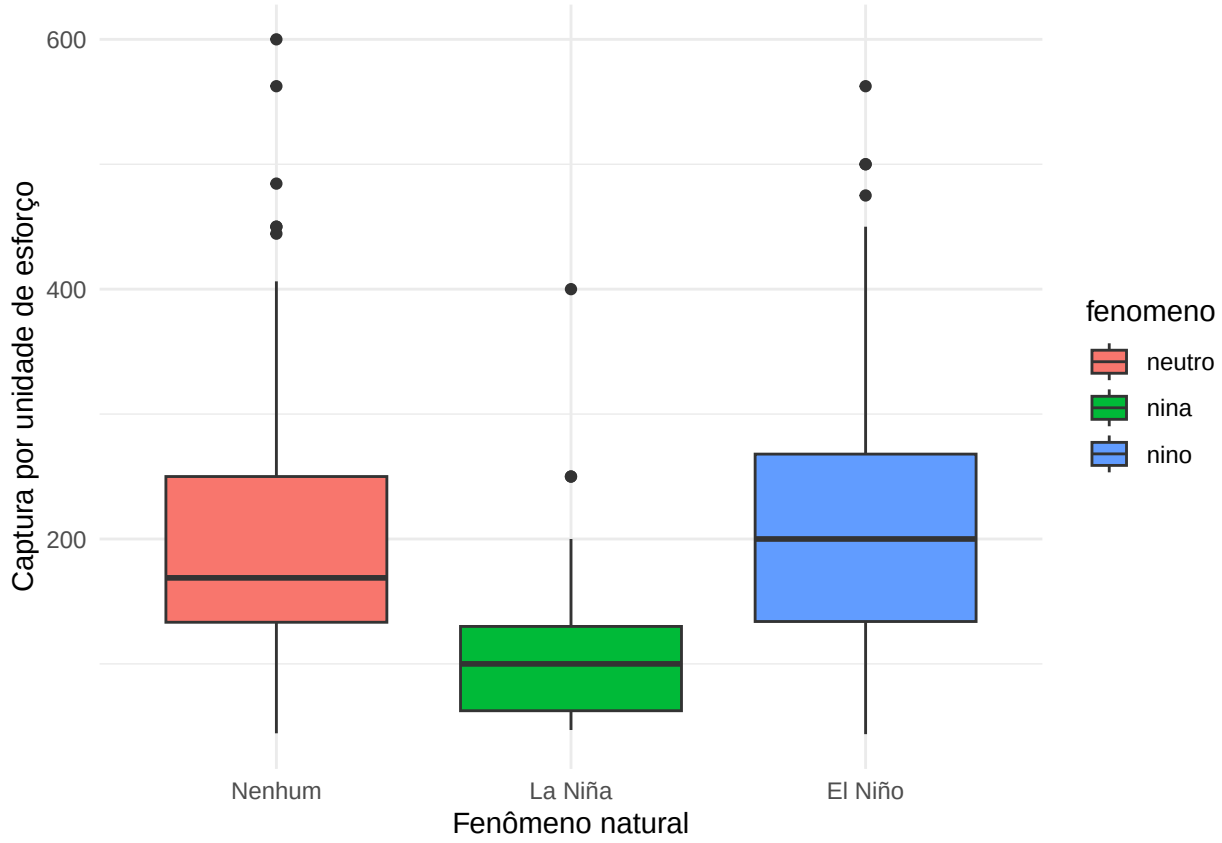
Análise dos Dados

```
##     frota                ano         trimestre        latitude
##  Length:156         Min.   :1995   Min.   :1.000   Min.   :23.25
##  Class :character   1st Qu.:1996   1st Qu.:2.000   1st Qu.:25.25
##  Mode  :character   Median :1998   Median :3.000   Median :26.25
##                     Mean   :1998   Mean   :2.679   Mean   :26.22
##                     3rd Qu.:1999   3rd Qu.:4.000   3rd Qu.:27.25
##                     Max.   :1999   Max.   :4.000   Max.   :28.25
##    longitude          cpue          fenomeno             cpue2
##  Min.   :41.25   Min.   : 43.75   Length:156         Min.   :3.778
##  1st Qu.:46.25   1st Qu.:108.33   Class :character   1st Qu.:4.684
##  Median :46.25   Median :166.41   Mode  :character   Median :5.114
##  Mean   :46.28   Mean   :195.55                      Mean   :5.086
##  3rd Qu.:46.75   3rd Qu.:250.00                      3rd Qu.:5.521
##  Max.   :48.25   Max.   :600.00                      Max.   :6.397
```

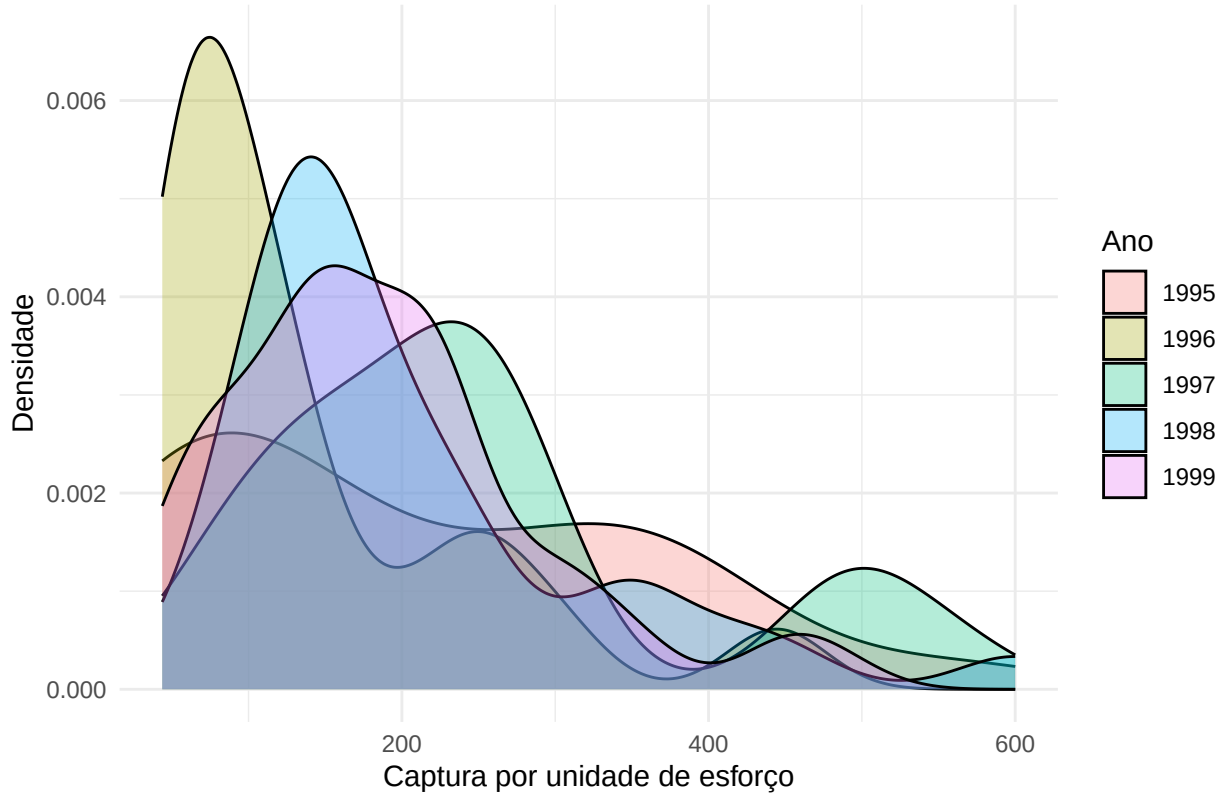
```
ggplot(dados2, aes(x = cpue)) +
  geom_density(alpha = 0.5, fill = "violet", col = "violet") +
  labs(
    #title = "Histograma de Densidade",
    x = "Captura por unidade de pesca",
    y = "Densidade"
  ) + theme_minimal() +
  geom_vline(xintercept = mean(dados2$cpue), linetype = "dashed", color = "black")
```



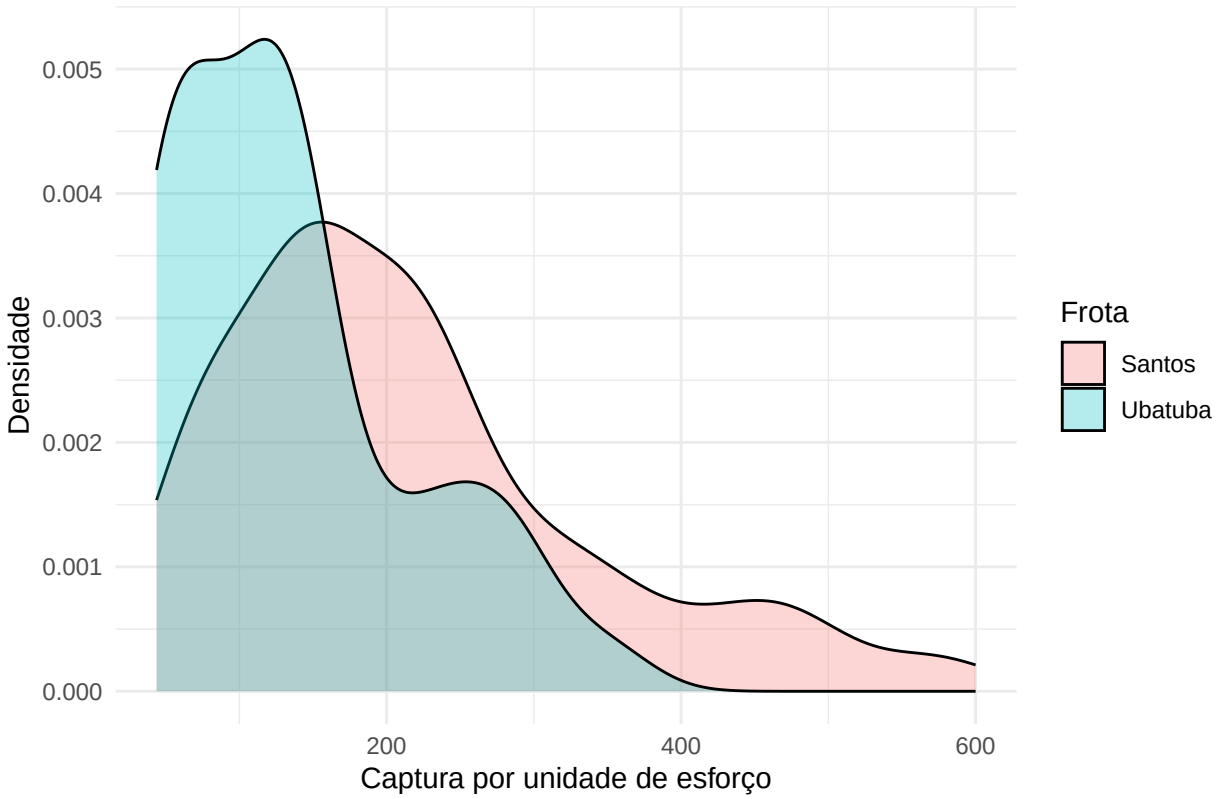

```
ggplot(dados2, aes(x = fenomeno, y = cpue, fill = fenomeno)) +
  geom_boxplot() +
  labs(
   # title = " ",
    x = "Fenômeno natural",
    y = "Captura por unidade de esforço"
  ) + scale_x_discrete(labels = c("Nenhum", "La Niña", "El Niño")) +
  theme_minimal()
```



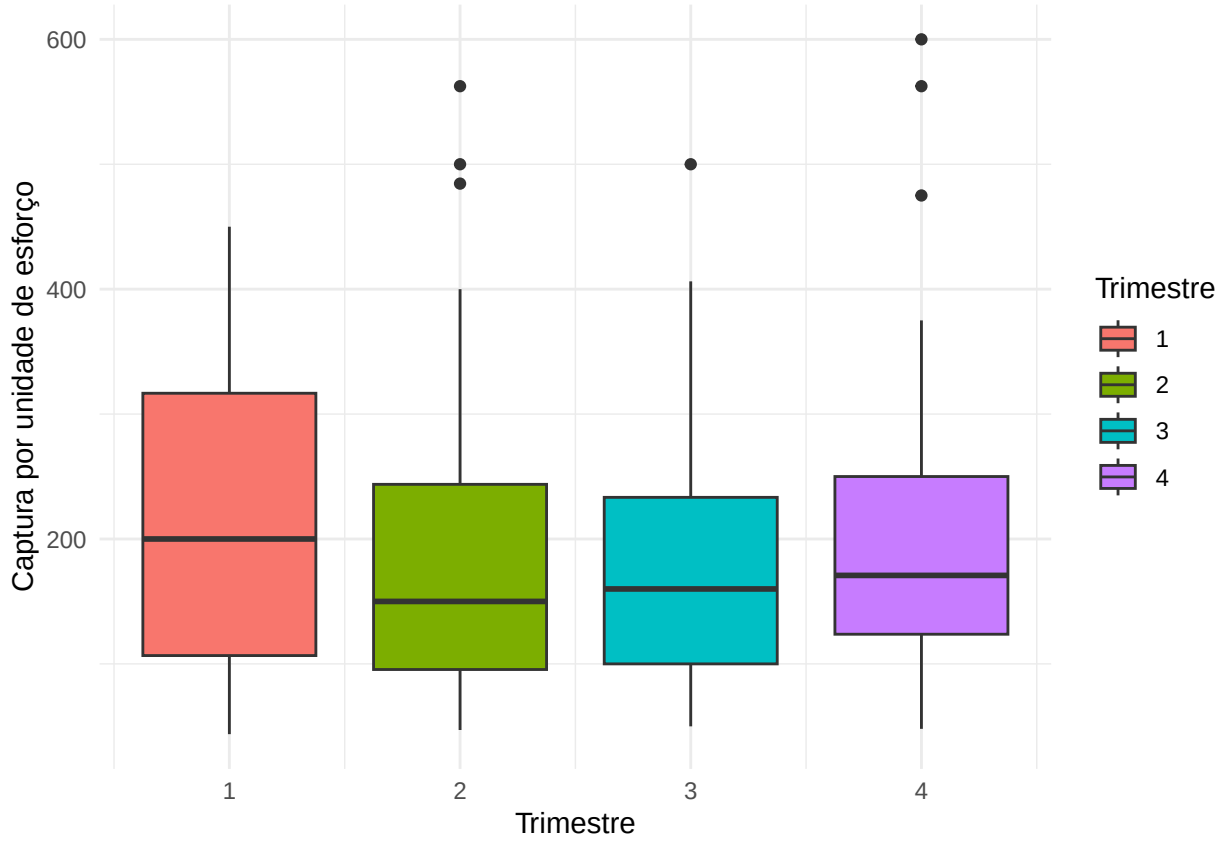

```
ggplot(dados2, aes(x = cpue, fill = as.factor(ano))) +
  geom_density(alpha = 0.3) +
  labs(
    title = " ",
    x = "Captura por unidade de esforço",
    y = "Densidade",
    fill = 'Ano'
  ) + theme_minimal()
```



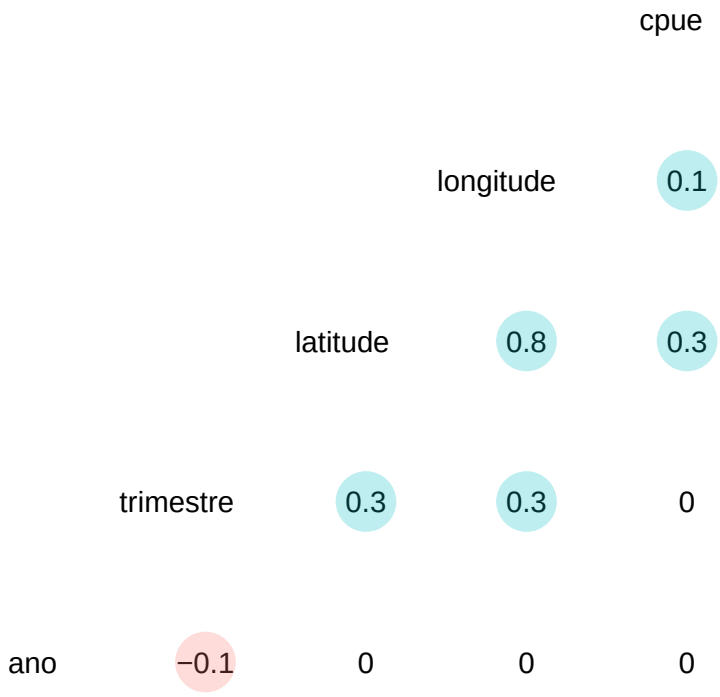

```
ggplot(dados2, aes(x = cpue, fill = frota)) +
  geom_density(alpha = 0.3) +
  labs(
    title = " ",
    x = "Captura por unidade de esforço",
    y = "Densidade",
    fill = 'Frota'
  ) + theme_minimal()
```




```
ggplot(dados2, aes(x = trimestre, y = cpue, fill = as.factor(trimestre))) +
  geom_boxplot() +
  labs(
    # title = " ",
    x = "Trimestre",
    y = "Captura por unidade de esforço",
    fill = "Trimestre"
  ) +
  theme_minimal()
```




```
ggcorr(select(dados2, -c(cpue2)), geom = "blank", label = TRUE, hjust = 0.75) +
  geom_point(size = 10, aes(color = coefficient >= 0, alpha = abs(coefficient) >= 0.05)) +
  scale_alpha_manual(values = c("TRUE" = 0.25, "FALSE" = 0)) +
  guides(color = FALSE, alpha = FALSE)
```




```
dados2 <- dados2 |> select(-c(cpue))

modelo <- stats::lm(cpue2 ~ ., data=dados2)
summary(modelo)
```

```
##
## Call:
## stats::lm(formula = cpue2 ~ ., data = dados2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.32394 -0.34570  0.02379  0.40315  1.23899
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)  32.8645028 83.3462696   0.394  0.69392
## frotaUbatuba -0.1852167  0.1288770  -1.437  0.15278
## ano          -0.0137336  0.0417482  -0.329  0.74265
## trimestre    -0.0006775  0.0463110  -0.015  0.98835
## latitude      0.2050073  0.0708918   2.892  0.00441 **
## longitude    -0.1213854  0.0744513  -1.630  0.10514
## fenomenonina -0.4816618  0.1720236  -2.800  0.00579 **
## fenomenonino  0.0828306  0.1220904   0.678  0.49855
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 0.5709 on 148 degrees of freedom
## Multiple R-squared:  0.2334, Adjusted R-squared:  0.1971
## F-statistic: 6.436 on 7 and 148 DF,  p-value: 1.27e-06
```

```
opt_model_step_aic<- stepAIC(modelo, direction="both")
```

```
## Start:  AIC=-167.11
## cpue2 ~ frota + ano + trimestre + latitude + longitude + fenomeno
##
##             Df Sum of Sq    RSS     AIC
## - trimestre  1    0.0001 48.234 -169.11
## - ano        1    0.0353 48.269 -169.00
## <none>                   48.234 -167.11
## - frota      1    0.6731 48.907 -166.95
## - longitude  1    0.8663 49.100 -166.34
## - latitude   1    2.7254 50.959 -160.54
## - fenomeno   2    4.1301 52.364 -158.30
##
## Step:  AIC=-169.11
## cpue2 ~ frota + ano + latitude + longitude + fenomeno
##
##             Df Sum of Sq    RSS     AIC
## - ano        1    0.0352 48.269 -171.00
## <none>                   48.234 -169.11
## - frota      1    0.6731 48.907 -168.95
## - longitude  1    0.8742 49.108 -168.31
## + trimestre  1    0.0001 48.234 -167.11
## - latitude   1    2.7427 50.976 -162.49
## - fenomeno   2    4.2000 52.434 -160.09
##
## Step:  AIC=-171
## cpue2 ~ frota + latitude + longitude + fenomeno
##
##             Df Sum of Sq    RSS     AIC
## <none>                   48.269 -171.00
## - frota      1    0.6794 48.948 -170.82
## - longitude  1    0.9236 49.193 -170.04
## + ano        1    0.0352 48.234 -169.11
## + trimestre  1    0.0001 48.269 -169.00
## - latitude   1    2.9133 51.182 -163.86
## - fenomeno   2    4.8533 53.122 -160.05
```

```
summary(opt_model_step_aic)
```

```
##
## Call:
## stats::lm(formula = cpue2 ~ frota + latitude + longitude + fenomeno,
##     data = dados2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2923 -0.3436  0.0156  0.4142  1.2353
```

```
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   5.45279    2.22339   2.452  0.01534 *
## frotaUbatuba -0.18604    0.12804  -1.453  0.14829
## latitude      0.20857    0.06932   3.009  0.00308 **
## longitude    -0.12413    0.07327  -1.694  0.09230 .
## fenomenonina -0.44620    0.13319  -3.350  0.00102 **
## fenomenonino  0.09949    0.10865   0.916  0.36130
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5673 on 150 degrees of freedom
## Multiple R-squared:  0.2328, Adjusted R-squared:  0.2072
## F-statistic: 9.104 on 5 and 150 DF,  p-value: 1.398e-07
```